



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/899,917

DATE: 01/24/2002

TIME: 19:23:38

Input Set : N:\Crf3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Olsen, Henrik S.

6 Ruben, Steven M.

8 (ii) TITLE OF INVENTION: Human Oncogene Induced Secreted Protein I

10 (iii) NUMBER OF SEQUENCES: 17

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Sterne, Kessler Goldstein & Fox P.L.L.C.

14 (B) STREET: 1100 New York Ave., NW, Suite 600

15 (C) CITY: Washington

16 (D) STATE: DC

17 (E) COUNTRY: USA

18 (F) ZIP: 20005-3934

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/899,917

C--> 28 (B) FILING DATE: 09-Jul-2001

39 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 08/994,962

33 (B) FILING DATE:

37 (A) APPLICATION NUMBER: US 60/037,388

38 (B) FILING DATE: 07-FEB-1997

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Steffe, Eric K.

43 (B) REGISTRATION NUMBER: 36,688

44 (C) REFERENCE/DOCKET NUMBER: 1488.0440002

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: 202-371-2600

48 (B) TELEFAX: 202-371-2540

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 860 base pairs

55 (B) TYPE: nucleic acid

56 (C) STRANDEDNESS: double

57 (D) TOPOLOGY: linear

59 (ii) MOLECULE TYPE: cDNA

62 (ix) FEATURE:

63 (A) NAME/KEY: CDS

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/899,917

DATE: 01/24/2002

TIME: 19:23:38

Input Set : N:\Crif3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

```

64      (B) LOCATION: 20..505
66      (ix) FEATURE:
67          (A) NAME/KEY: sig_peptide
68          (B) LOCATION: 20..79
70      (ix) FEATURE:
71          (A) NAME/KEY: mat_peptide
72          (B) LOCATION: 80..505
75      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
77 TCCCATACAG GCCCCCACC ATG AAG GGT TTC ACA GCC ACT CTC TTC CTC TGG      52
78          Met Lys Gly Phe Thr Ala Thr Leu Phe Leu Trp
79          -20                      -15                      -10
81 ACT CTG ATT TTT CCC AGC TGC AGT GGA GGC GGC GGT GGG AAA GCC TGG      100
82 Thr Leu Ile Phe Pro Ser Cys Ser Gly Gly Gly Gly Lys Ala Trp
83          -5                      1                      5
85 CCC ACA CAC GTG GTC TGT AGC GAC AGC GGC TTG GAA GTG CTC TAC CAG      148
86 Pro Thr His Val Val Cys Ser Asp Ser Gly Leu Glu Val Leu Tyr Gln
87          10                      15                      20
89 AGT TGC GAT CCA TTA CAA GAT TTT GGC TTT TCT GTT GAA AAG TGT TCC      196
90 Ser Cys Asp Pro Leu Gln Asp Phe Gly Phe Ser Val Glu Lys Cys Ser
91          25                      30                      35
93 AAG CAA TTA AAA TCA AAT ATC AAC ATT AGA TTT GGA ATT ATT CTG AGA      244
94 Lys Gln Leu Lys Ser Asn Ile Asn Ile Arg Phe Gly Ile Ile Leu Arg
95          40                      45                      50                      55
97 GAG GAC ATC AAA GAG CTT TTT CTT GAC CTA GCT CTC ATG TCT CAA GGC      292
98 Glu Asp Ile Lys Glu Leu Phe Leu Asp Leu Ala Leu Met Ser Gln Gly
99          60                      65                      70
101 TCA TCT GTT TTG AAT TTC TCC TAT CCC ATC TGT GAG GCG GCT CTG CCC      340
102 Ser Ser Val Leu Asn Phe Ser Tyr Pro Ile Cys Glu Ala Ala Leu Pro
103          75                      80                      85
105 AAG TTT TCT TTC TGT GGA AGA AGG AAA GGA GAG CAG ATT TAC TAT GCT      388
106 Lys Phe Ser Phe Cys Gly Arg Arg Lys Gly Glu Gln Ile Tyr Tyr Ala
107          90                      95                      100
109 GGG CCT GTC AAT AAT CCT GAA TTT ACT ATT CCT CAG GGA GAA TAC CAG      436
110 Gly Pro Val Asn Asn Pro Glu Phe Thr Ile Pro Gln Gly Glu Tyr Gln
111          105                      110                      115
113 GTT TTG CTG GAA CTG TAC ACT GAA AAA CGG TCC ACC GTG GCC TGT GCC      484
114 Val Leu Leu Glu Leu Tyr Thr Glu Lys Arg Ser Thr Val Ala Cys Ala
115          120                      125                      130                      135
117 AAT GCT ACT ATC ATG TGC TCC TGA CTGTGGC CTGTAGCAAA AATCACAGCC      535
118 Asn Ala Thr Ile Met Cys Ser
119          140
121 AGCTGCATCT CGTGGGACCT CCAAGCTCCT CTGACTGAAC CTACTGTGGG AGGAGAAGCA      595
123 GCTGATGACA GAGAGAGGCT CTACAAAGAA GCGCCCCCAA AGAGTGCAGC TGCTAATTTT      655
125 AGTCCCAGGA CCAGACATCC CCAGACTCCA CAGATGTAAT GAAGTCCCGG AATGTATCTG      715
127 TTTCTAAGGA GCCTCTTGGC AGTCCTTAAG CAGTCTTGAG GGTCCATCCT TTTTCTCTAA      775
129 TTGGTCGCCT CCCACCAGAC TCACCTGCTT TTCAACTTTT TAGGAGTGCT TCCTCACAGT      835
131 TACCAAGAAA TAAAGAAAGC TGGCC      860
134 (2) INFORMATION FOR SEQ ID NO: 2:
136      (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/899,917

TIME: 19:23:38

Input Set : N:\Crif3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

```

137      (A) LENGTH: 162 amino acids
138      (B) TYPE: amino acid
139      (D) TOPOLOGY: linear
141      (ii) MOLECULE TYPE: protein
143      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
145 Met Lys Gly Phe Thr Ala Thr Leu Phe Leu Trp Thr Leu Ile Phe Pro
146 -20      -15      -10      -5
148 Ser Cys Ser Gly Gly Gly Gly Lys Ala Trp Pro Thr His Val Val
149      1      5      10
151 Cys Ser Asp Ser Gly Leu Glu Val Leu Tyr Gln Ser Cys Asp Pro Leu
152      15      20      25
154 Gln Asp Phe Gly Phe Ser Val Glu Lys Cys Ser Lys Gln Leu Lys Ser
155      30      35      40
157 Asn Ile Asn Ile Arg Phe Gly Ile Ile Leu Arg Glu Asp Ile Lys Glu
158 45      50      55      60
160 Leu Phe Leu Asp Leu Ala Leu Met Ser Gln Gly Ser Ser Val Leu Asn
161      65      70      75
163 Phe Ser Tyr Pro Ile Cys Glu Ala Ala Leu Pro Lys Phe Ser Phe Cys
164      80      85      90
166 Gly Arg Arg Lys Gly Glu Gln Ile Tyr Tyr Ala Gly Pro Val Asn Asn
167      95      100      105
169 Pro Glu Phe Thr Ile Pro Gln Gly Glu Tyr Gln Val Leu Leu Glu Leu
170      110      115      120
172 Tyr Thr Glu Lys Arg Ser Thr Val Ala Cys Ala Asn Ala Thr Ile Met
173 125      130      135      140
175 Cys Ser
178 (2) INFORMATION FOR SEQ ID NO: 3:
180      (i) SEQUENCE CHARACTERISTICS:
181      (A) LENGTH: 133 amino acids
182      (B) TYPE: amino acid
183      (C) STRANDEDNESS: Not Relevant
W--> 184      (D) TOPOLOGY: Not Relevant
186      (ii) MOLECULE TYPE: protein
191      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
193 Trp Pro Thr His Thr Val Cys Lys Glu Glu Asn Leu Glu Ile Tyr Tyr
194 1      5      10      15
196 Lys Ser Cys Asp Pro Gln Gln Asp Phe Ala Phe Ser Ile Asp Arg Cys
197      20      25      30
199 Ser Asp Val Thr Thr His Thr Phe Asp Ile Arg Ala Ala Met Val Leu
200      35      40      45
202 Arg Gln Ser Ile Lys Glu Leu Tyr Ala Lys Val Asp Leu Ile Ile Asn
203      50      55      60
205 Gly Lys Thr Val Leu Ser Tyr Ser Glu Thr Leu Cys Gly Pro Gly Leu
206      65      70      75      80
208 Ser Lys Leu Ile Phe Cys Gly Lys Lys Lys Gly Glu His Leu Tyr Tyr
209      85      90      95
211 Glu Gly Pro Ile Thr Leu Gly Ile Lys Glu Ile Pro Gln Gly Asp Tyr
212      100      105      110
214 Thr Ile Thr Ala Arg Leu Thr Asn Glu Asp Arg Ala Thr Val Ala Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/899,917

DATE: 01/24/2002

TIME: 19:23:38

Input Set : N:\Crf3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

```

215          115          120          125
217      Ala Asp Phe Thr Val
218          130
220 (2) INFORMATION FOR SEQ ID NO: 4:
222     (i) SEQUENCE CHARACTERISTICS:
223         (A) LENGTH: 29 base pairs
224         (B) TYPE: nucleic acid
225         (C) STRANDEDNESS: single
226         (D) TOPOLOGY: linear
228     (ii) MOLECULE TYPE: cDNA
233     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
235 GACTCCATGG GCGGCGGTGG GAAAGCCTG                29
237 (2) INFORMATION FOR SEQ ID NO: 5:
239     (i) SEQUENCE CHARACTERISTICS:
240         (A) LENGTH: 30 base pairs
241         (B) TYPE: nucleic acid
242         (C) STRANDEDNESS: single
243         (D) TOPOLOGY: linear
245     (ii) MOLECULE TYPE: cDNA
250     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
252 GACTAGATCT GGAGCACATG ATAGTAGCAT                30
254 (2) INFORMATION FOR SEQ ID NO: 6:
256     (i) SEQUENCE CHARACTERISTICS:
257         (A) LENGTH: 36 base pairs
258         (B) TYPE: nucleic acid
259         (C) STRANDEDNESS: single
260         (D) TOPOLOGY: linear
262     (ii) MOLECULE TYPE: cDNA
267     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
269 GACTGGATCC GCCATCATGA AGGGTTTCAC AGCCAC        36
271 (2) INFORMATION FOR SEQ ID NO: 7:
273     (i) SEQUENCE CHARACTERISTICS:
274         (A) LENGTH: 29 base pairs
275         (B) TYPE: nucleic acid
276         (C) STRANDEDNESS: single
277         (D) TOPOLOGY: linear
279     (ii) MOLECULE TYPE: cDNA
284     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
286 GACTGGTACC AGCAGCTGCA CTCTTTGGG                29
288 (2) INFORMATION FOR SEQ ID NO: 8:
290     (i) SEQUENCE CHARACTERISTICS:
291         (A) LENGTH: 36 base pairs
292         (B) TYPE: nucleic acid
293         (C) STRANDEDNESS: single
294         (D) TOPOLOGY: linear
296     (ii) MOLECULE TYPE: cDNA
301     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
303 AGCTAAGCTT CCGCCACCAT GAAGGGTTTC ACAGCC        36
305 (2) INFORMATION FOR SEQ ID NO: 9:

```

RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/899,917

TIME: 19:23:38

Input Set : N:\Crf3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

```

307      (i) SEQUENCE CHARACTERISTICS:
308          (A) LENGTH: 62 base pairs
309          (B) TYPE: nucleic acid
310          (C) STRANDEDNESS: single
311          (D) TOPOLOGY: linear
313      (ii) MOLECULE TYPE: cDNA
318      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
320 CAGTCTCGAG TTAAGCGTAG TCTGGGACGT CGTATGGGTA GGAGCACATG ATAGTAGCAT      60
322 TG                                                                    62
324 (2) INFORMATION FOR SEQ ID NO: 10:
326      (i) SEQUENCE CHARACTERISTICS:
327          (A) LENGTH: 29 base pairs
328          (B) TYPE: nucleic acid
329          (C) STRANDEDNESS: single
330          (D) TOPOLOGY: linear
332      (ii) MOLECULE TYPE: cDNA
337      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
339 GACTGGTACC AGCAGCTGCA CTCTTTGGG      29
341 (2) INFORMATION FOR SEQ ID NO: 11:
343      (i) SEQUENCE CHARACTERISTICS:
344          (A) LENGTH: 514 base pairs
345          (B) TYPE: nucleic acid
346          (C) STRANDEDNESS: double
347          (D) TOPOLOGY: linear
349      (ii) MOLECULE TYPE: cDNA
354      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
356 NAATTCGCGA GATTTTCCC AGCTGCAAGT GAGGCGGCGG TGGGAAAGCC TGGCCACAC      60
358 ACGTGGTCTG TAGCGACAGG CTTTGGAGGT GCTCTACCAG AGTTGCGATC CATTACAAGA      120
360 TTTTGGCTTT TCTGTTGAAA AGTGTTCCAA GCAATTAATA TCAAATATCA ACATTAGATT      180
362 TGGAATTATT CTGAAGGACA TCAAAGAGCT TTTTCTTGAC CTAGCTCTCA TGTNTCAAGG      240
364 CTCATCTGTT TTGAATTTCT CCTATCCCAT CTGTGAGGCG GCTCTGCCAA GTTTTCTTTC      300
366 TGTGGAAGAA GGAAAGGAGA GCAGATTTAC TATGCTNGGG CTGTCAATAA TNCNGAATTT      360
368 ACTATTTCCT CANGGGGGAT TACCAGGTTT TGCTGGGACT GTACAATGAA AAACGGTCCA      420
370 CCGNGGCNGT GCCATGGTAC TATCGNGTGG TCCGACTGTG GCCNTAGGAA AATCACACCA      480
372 TTGNATTCGG GGNCNCCAGT CCTTGATNAC CNAN      514
374 (2) INFORMATION FOR SEQ ID NO: 12:
376      (i) SEQUENCE CHARACTERISTICS:
377          (A) LENGTH: 457 base pairs
378          (B) TYPE: nucleic acid
379          (C) STRANDEDNESS: double
380          (D) TOPOLOGY: linear
382      (ii) MOLECULE TYPE: cDNA
387      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
389 CACAGCCACT CTCTTCTCTT GGACTCTAAT TTTNCCCAGC TGCAGTGGAG GCGGCGGTGG      60
391 GAAAGCCTGG CCCACACACG TGGTCTGTAG CGACANGGCT TGGAAGTGCT CTACCAGAGT      120
393 TGCGATCCAT TACAAGATTT TGGCTTTTCT GTTGAAAAGT GTTCCAAGCA ATTAATAATCA      180
395 AATATCAACA TTAGATTTGG AATTATTCTG AGAGAGGACA TCAAAGAGCT TTTTCTTGAC      240
397 CTAGCTCTCA TGTCTCAAGG CTCATCTGTT TTNAATTTCT CCTATCCCAT CTGTNAGGCG      300
399 GCTCTGCCCA AGTTTTCTTT CTGTGGAAGA AGGAAAGGAG AGCAGATTTA CTATGCTGGG      360

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/899,917

DATE: 01/24/2002

TIME: 19:23:39

Input Set : N:\Crf3\RULE60\09899917.raw

Output Set: N:\CRF3\01242002\I899917.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:184 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=3
L:516 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=17